Slow Manifolds and Invariant Sets of the Primitive Equations

R. TEMAM

Institute for Scientific Computing and Applied Mathematics, Indiana University, Bloomington, Indiana

D. WIROSOETISNO

Mathematical Sciences, Durham University, Durham, United Kingdom

(Manuscript received 20 August 2010, in final form 2 November 2010)

ABSTRACT

The authors review, in a geophysical setting, several recent mathematical results on the forced–dissipative hydrostatic primitive equations with a linear equation of state in the limit of strong rotation and stratification, starting with existence and regularity (smoothness) results and describing their implications for the long-time behavior of the solution. These results are used to show how the solution of the primitive equations in a periodic box comes close to geostrophic balance as $t \rightarrow \infty$. Then a review follows of how geostrophic balance could be extended to higher orders in the Rossby number, and it is shown that the solution of the primitive equations also satisfies a higher-order balance up to an exponentially small error. Finally, the connection between balance dynamics in the primitive equations and its global attractor, which is the only known invariant set (for a sufficiently general forcing), is discussed.

1. Introduction

Ever since computers became fast enough to integrate the full primitive equations (PEs), efforts have been made to filter fast oscillations (''gravity waves'') from the solutions. Early works (e.g., Baer and Tribbia 1977; Machenhauer 1977) sought to bound fast terms in the time derivatives at time $t = 0$, but it was found that no matter how carefully this initialization was done, rapid oscillations eventually developed (see Daley 1991, section 6.7 for a review of the early efforts), although this may take a long time even for a reasonably large model (Errico 1984; see also Vautard and Legras 1986).

Leith (1980) and Lorenz (1980) introduced the concept of a slow manifold devoid of fast oscillations, where slow dynamics, if it exists, is hypothesized to take place. Soon thereafter, however, Warn (1997)[1] argued that the existence of an invariant manifold devoid of fast oscillations is an exception rather than the rule, and that the proposed constructions are likely to be asymptotic rather than convergent. Using special structures of a simple model, Lorenz (1986) did succeed in finding manifolds that are invariant, but these are nearly discontinuous with respect to the variables and still contain a small amount of fast oscillations.

In light of this early evidence, Warn (1997), Warn and Ménard (1986), and Lorenz (1986) proposed, in place of a slow and invariant manifold, a thin layer (which they termed a ''fuzzy manifold'') in which gravity wave activity is minimal. It was understood, sometimes implicitly, that viscosity or gravity wave radiation would be needed for the solution to arrive at, and stay within, this layer. Numerical evidence from more realistic models (e.g., Mohebalhojeh and McIntyre 2007, and references therein) suggests that balance can be very accurate; that is, this layer can be very thin, but never exact (at least for nonsteady flows). The discovery by Vanneste and Yavneh (2004) of an exact solution of the inviscid Boussinesq primitive equation that supports spontaneous generation of gravity waves effectively ruled out the existence of an invariant slow manifold in general. The amplitude of gravity waves thus generated scales as $\exp(-c/\varepsilon)$, so they cannot be captured by formal perturbative approaches, but at the same time it explains why the latter often work very well in practice.

[1] This insightful paper was written in 1983, circulated and cited widely, but not published (in its original form) until 1997.

Corresponding author address: D. Wirosoetisno, Mathematical Sciences, Durham University, Durham DH1 3LE, United Kingdom.
E-mail: djoko.wirosoetisno@durham.ac.uk

DOI: 10.1175/2010JAS3650.1

In this paper we review several recent mathematical results on the primitive equations that prove rigorously that, under some reasonable hypotheses, the solution will enter an exponentially thin layer around a manifold on which fast oscillations are exponentially weak. The main mechanism is direct damping by viscosity, which also ensures the existence and smoothness of the solution over the long times needed to achieve balance. The more physically appealing mechanisms such as gravity wave radiation and turbulent cascade pose an altogether different, and challenging, mathematical problem, so we shall relegate it to the discussion (and future works).

Unless necessary for our discussion, we will be deliberately informal when it comes to norms, function spaces, constants, etc. Readers who are interested in the details should consult the mathematical works cited below for precise statements of the results.

2. Global bounds and attractors

We consider a fluid with a linear equation of state $\rho = \rho_0 - \alpha T + \beta s$, where T is the temperature and s denotes other factors such as salinity. If the diffusivities for T and s are equal, which will be assumed henceforth for simplicity, their effect in the dynamics can be represented simply by the density ρ, or equivalently the buoyancy b. Now assuming that the fluid is stably stratified, $\rho(x, y, z, t) = \rho_0 - z\rho_1 + \rho'(x, y, z, t)$, where ρ_0 and ρ_1 are positive constants, we impose the Boussinesq approximation and hydrostatic balance.

For concreteness, we fix the length scale L as the horizontal size of the periodic box and some arbitrary velocity scale V. Let f be the Coriolis parameter and $N = \sqrt{g\rho_1/\rho_0}$ the Brunt–Väisälä frequency. In this article, we consider the limit of strong rotation and stratification by taking $\varepsilon := V/(fL) = V/(NH)$ small. (If one thought of V as the typical velocity scale, then ε would be the classical Rossby and Froude numbers, but, as discussed below, these numbers may not give enough information on the dynamics, even in some asymptotic limit.) With this scaling, the nondimensional primitive equations read

$$
\begin{aligned}
\partial_t \mathbf{v} + \frac{1}{\varepsilon}[\mathbf{v}^\perp + \nabla p] + \mathbf{v}\cdot\nabla\mathbf{v} + w\partial_z\mathbf{v} &= \mu\Delta_3\mathbf{v} + \mathbf{S}_\mathbf{v},\\
\partial_t b + \frac{1}{\varepsilon}w + \mathbf{v}\cdot\nabla b + w\partial_z b &= \mu\Delta_3 b + S_b,\\
\nabla\cdot\mathbf{v} + \partial_z w &= 0, \quad \text{and}\\
b &= \partial_z p. \qquad (1)
\end{aligned}
$$

Here $\mathbf{v} := (u, v)$ is the horizontal velocity and $\mathbf{v}^\perp := (-v, u)$ denotes the Coriolis term; b is the buoyancy, $\Delta_3 := \partial_{xx} + \partial_{yy} + \partial_{zz}$: that is, both $\mathbf{v}$ and b are diffused in all three directions. Mathematically, this is needed for the regularity of the solution and can be regarded as arising from the diffusions of momentum and heat. The diffusion constants μ are (scaled) inverse Reynolds numbers, which for simplicity of analysis have been set equal. The sources of momentum $\mathbf{S}_\mathbf{v}$ and of buoyancy S_b ensure that the long-term dynamics is nontrivial. The regularity and leading-order geostrophic decay results (sections 2 and 3) hold for time-dependent $\mathbf{S}$, but the existence of the global attractor and of higher-order balance (section 4) require $\mathbf{S}$ to be time independent (or, with some extra work, at most quasiperiodic).

For boundary conditions, we assume periodicity in the horizontal directions and free-slip $\partial_z\mathbf{v} = 0$, $w = 0$, and $b = 0$, at the bottom and top. Spectrally, this can be implemented by doubling the domain in the z direction and imposing appropriate symmetries (see, e.g., Bartello 1995). These boundary conditions are not needed for the regularity of the solutions [note that Cao and Titi (2007), Kobelkov (2007), and Kukavica and Ziane (2007) used different boundary conditions], but they are essential for the existence of our exponential slow manifold. With no loss of generality, we also assume that $\boldsymbol{v}$ has zero integral over the domain.

Taking the scalar product of (1a) with $\mathbf{v}$ and (1b) with b and integrating by parts (i.e., using conservation of kinetic energy and of buoyancy), the conservative terms all cancel. Including the forcing and dissipative terms, we find, upon using Poincaré and Cauchy–Schwarz inequalities, the differential inequality

$$
\frac{d}{dt}\int \{|\mathbf{v}|^2 + b^2\}\, d\mathbf{x}^3 + \mu\int\{|\nabla_3\mathbf{v}|^2 + |\nabla_3 b|^2\}\, d\mathbf{x}^3 \le \frac{c_1}{\mu}\int\{|\mathbf{S}_\mathbf{v}| + |S_b|\}\, d\mathbf{x}^3. \qquad (2)
$$

Here and in what follows, c_i are constants that depend on the size of the domain but on no other parameter. If the forcing $\mathbf{S}$ depends on time, the integral on the right-hand side (rhs) of (2) and (4) below is to be replaced by its maximum over $t \ge 0$. In the absence of forcing and dissipation, the first integral in (2) is clearly constant in time.

Let $U := (\mathbf{v}, b)$ and denote

$$
E_0(U) := \frac{1}{2}\int\{|\mathbf{v}(\mathbf{x})|^2 + b(\mathbf{x})^2\}\, d\mathbf{x}^3. \qquad (3)
$$

Integrating the differential inequality (2), we find that $E_0(U(t))$ is bounded uniformly for all time:

$$E_0(U(t)) \le e^{-c\mu t}E_0(U(0)) + (1-e^{-c\mu t})\frac{c_2}{\mu}\int\{|\mathbf{S_v}|^2 + |S_b|^2\}\,d\mathbf{x}^3. \qquad (4)$$

Physically, this simply says that the influence of the initial conditions decays in time and the solution is bounded more and more by the forcing. It is clear that we can find a time $T_0((\mathbf{S}, U(0); \mu)$, where $\mathbf{S} = (\mathbf{S_v}, S_b)$, such that

$$E_0(U(t)) \le \frac{2c_2}{\mu}\int\{|\mathbf{S_v}|^2 + |S_b|^2\}\,d\mathbf{x}^3 =: K_0(\mathbf{S};\mu) \qquad (5)$$

for all $t \ge T_0$. Similarly to the 3D Navier–Stokes equations, however, the fact that E_0 is bounded uniformly in time is not enough to ensure that the solution remains smooth (even if the initial conditions and forcing are) and unique for all time. For existence and uniqueness for all time, we also require that

$$E_1(U) := \frac{1}{2}\int\{|\nabla_3\mathbf{v}(\mathbf{x})|^2 + |\nabla_3 b(\mathbf{x})|^2\}\,d\mathbf{x}^3 \qquad (6)$$

be finite for all finite time. (In fact, since $\mathbf{v}$ has zero average and b vanishes at the top and bottom, the fact that E_1 is bounded also implies the boundedness of E_0.)

In the absence of mean stratification [i.e., without the w/ε term in (1b)], this problem was finally solved in 2005: Cao and Titi (2007) and Kobelkov (2007) independently proven that $E_1(U(t)) < \infty$ for all t if it holds at $t = 0$ (and if $\mathbf{S}$ is suitably bounded).[2] The key ingredient in their proofs is hydrostatic balance:[3] it has a regularizing effect on the solution. Further refinement by Ju (2007) gives us a uniform bound,[4] $E_1(U(t)) \le \tilde{N}_1(\mathbf{S}, U(0); \mu, \varepsilon)$, valid for all $t \ge 0$.

When the forcing $\mathbf{S}$ does not depend on time, following works for the Navier–Stokes equations (see, e.g., Temam 1997), Ju proved that this implies the existence of the global attractor for the primitive equations. For our purposes here, the global attractor of a dynamical system (which may be infinite dimensional) is a compact set $\mathcal{A}$ in phase space with the following properties:

(A1) $\mathcal{A}$ is invariant[5] and is the largest such set; and
(A2) $\mathcal{A}$ attracts all solutions and is the smallest such set.

As Lorenz (1963) already realized, $\mathcal{A}$ appears[6] to be a fractal set with a very complicated structure even for simple systems, so the situation for partial differential equations is presumably worse. Moreover, despite much effort, no rigorous proof that $\mathcal{A}$ is continuous with respect to the parameters has been obtained. The presumed complicated structure (in phase and parameter spaces) of $\mathcal{A}$ makes it difficult to work with in our search for slow and/or invariant manifold—at this point in the argument, $\mathcal{A}$ may not be slow even at leading order. In other words, while an invariant set does exist for the primitive equations, it is probably neither a manifold nor slow.

Modifying the argument in Ju (2007), one can prove that the uniform bound $E_1(U(t)) \le \tilde{N}_1(\mathbf{S}, U(0); \mu, \varepsilon)$ also holds for the system (1). For the development in the next section, however, we need a stronger (as yet unproved) bound that is independent of ε,

$$E_1(U(t)) \le N_1(\mathbf{S}, U(0); \mu) \quad \text{for all} \quad t \ge 0. \qquad (7)$$

In what follows, we shall assume that (7) holds. In analogy with (4), one can then prove that for t sufficiently large, $E_1(U(t))$ is bounded independently of the initial conditions. Furthermore, assuming that the forcing (but not necessarily the initial conditions) is sufficiently smooth, one can prove (Petcu and Wirosoetisno 2005) that all derivatives of the solution are similarly bounded: with

$$E_n(U) := \frac{1}{2}\int\{|\nabla_3^n\mathbf{v}(\mathbf{x})|^2 + |\nabla_3^n b(\mathbf{x})|^2\}\,d\mathbf{x}^3, \qquad (8)$$

one has, for $n = 0, 1, 2, \ldots,$

$$E_n(U(t)) \le K_n(\mathbf{S}, \mu) \quad \text{for all} \quad t \ge T_n(U(0), \ldots). \qquad (9)$$

For our exponential estimate below, we need something stronger, resembling analyticity in finite-dimensional systems. Let $\mathbf{v_k}(t)$ and $b_\mathbf{k}(t)$ be the Fourier coefficients of $\mathbf{v}$ and b, and for $\sigma > 0$ define

$$E_1^\sigma(U) := \sum_\mathbf{k} e^{2\sigma|\mathbf{k}|}|\mathbf{k}|^2\{|\mathbf{v_k}|^2 + |b_\mathbf{k}|^2\}. \qquad (10)$$

Now if E_1^σ is finite, $\mathbf{v_k}$ and $b_\mathbf{k}$ must decay exponentially as $|\mathbf{k}| \to \infty$; such functions are said to have Gevrey regularity. Assuming that the forcing $\mathbf{S}$ is Gevrey, one can prove following Foias and Temam (1989) that the solution $(\boldsymbol{v}, b)$ will also be Gevrey (Petcu and Wirosoetisno 2005), in the sense that $E_1^\sigma(U(t))$ is bounded uniformly for all $t \ge 0$, and independently of the initial conditions

[2] But their bound on $E_1(U(t)) < \infty$ as $t \to \infty$; all that is needed for regularity is that $E_1(U(t))$ does not blow up at finite t.

[3] The nonhydrostatic case has been shown, for sufficiently strong rotation, to have a unique solution for all time by Babin et al. (2000).

[4] We stress that such bounds, having to account for all possible worst-case scenarios, usually do not give any meaningful estimate on how large the solution actually is; unhelpfully, "estimate" means "bound" in the mathematical literature.

[5] A subtler point: $\mathcal{A}$ is invariant both forward and backward in time even if a general solution is defined only for forward time.

[6] The complex structure of the Lorenz attractor was proved in 1999 by Tucker (2002).

$$E_1^\sigma(U(t)) \le K_1^\sigma(\mathbf{S};\, \mu) \tag{11}$$

for all $t > T_\sigma$. Informally, this means that the solution will become "analytic" after some time if the forcing is analytic, even if the initial conditions are not.

In the next sections, we describe how these regularity properties, far from being mere mathematical curiosities, can be used to obtain results on the balance behavior of the solution.

3. Decay to geostrophy

From the time of Charney (1948), the quasigeostrophic system

$$\partial_t q + \mathbf{v}_q \cdot \nabla q = \mu\Delta_3 q + f_q, \tag{12}$$

where $\mathbf{v}_q = \nabla^\perp \Delta_3^{-1} q$, has been used as an approximation to the primitive equations, (1) or its variants. When ε is small and the initial conditions for the full primitive equations (1) are at (or near) geostrophic balance, it has been proven that, subject to some smoothness conditions, the solution of (12) is a good approximation to the solution of (1) over a time of order 1; see Bourgeois and Beale (1994) and Babin et al. (2000) for the inviscid case in the Boussinesq PE, and Temam and Wirosoetisno (2007) for (1). Physically, this is not surprising, the more interesting question being whether and how the solution of (1) comes near geostrophic balance to begin with.

In Temam and Wirosoetisno (2010, hereafter TW10) we used the regularity results (9) to prove that, for any bounded initial conditions, the solution of the PE (1) will eventually end up near geostrophic balance. More precisely, let $q = \nabla^\perp \cdot \mathbf{v} + \partial_z b$ and $\psi = \Delta_3^{-1} q$ be the quasigeostrophic potential vorticity and streamfunction. We then split $(\mathbf{v}, b)$ into its geostrophic part $\mathbf{v}^0 = \nabla^\perp\psi$ and $b^0 = \partial_z\psi$, and ageostrophic part $\mathbf{v}^\varepsilon := \mathbf{v} - \mathbf{v}^0$ and $b^\varepsilon := b - b^0$. Assuming that the forcing is uniformly bounded,

$$\int \{|\nabla^2 \mathbf{S}(t)|^2 + |\partial_t \mathbf{S}(t)|^2\}\, d\mathbf{x}^3 \le K_f < \infty \tag{13}$$

for all $t \ge 0$, we have

$$\int \{|\mathbf{v}^\varepsilon(t)|^2 + |b^\varepsilon(t)|^2\}\, d\mathbf{x}^3 \le \varepsilon K_g(K_f, \mu) \tag{14}$$

for $t \ge T_g(\mathbf{v}(0), b(0), \mathbf{S}, \ldots)$. In other words, after a sufficiently long time, the ageostrophic energy of our solution will be $O(\varepsilon)$, whatever its initial value is.[7] We note that the forcing may be time dependent as long as its time derivative is bounded independently of ε; physically, this means that any "diurnal" forcing must be $O(\varepsilon)$ while "seasonal" forcing may be of order 1.

Since the attractor $\mathcal{A}$ attracts all solutions, our result implies that $\mathcal{A}$ must lie inside the set defined by (14); that is, as $\varepsilon \to 0$, the invariant set of the PE becomes closer and closer to being quasigeostrophic. Strictly speaking, this in itself does not prove that our solution is "slow"—it could execute small fast oscillations about geostrophic balance. Using higher-order results below, however, one can rule out this scenario.

Now let us visit one issue we have so far avoided: the Rossby number $\mathrm{Ro} := \tilde{V}/(f\tilde{L})$ where the Coriolis parameter f is given but the velocity gradient $\tilde{V}/\tilde{L}$ is now "typical" of the solution. On the right-hand side of (14), only ε explicitly contains f, while $K_g(\cdots)$ contains L^2 bounds on U, ∇U, $\nabla^2 U$, and $\nabla^3 U$. Thus, there is no simple "representative" $\tilde{V}/\tilde{L}$ in terms of which we can write the bound (14). Even worse, since K_g encodes long-term behavior of these velocity gradients, this bound is not local in time either. Not surprisingly, higher-order derivatives will figure in the higher-order balance in the next section. Now it is true that (14) is rather "pessimistic" since it has to cover all possible solutions of (1), and that in many situations (exact solutions, numerical simulations, etc.) one may be able to find a $\tilde{V}/(f\tilde{L})$ on which the observed imbalance appears to scale. We believe, however, that since balance is inherently nonlocal both in space and time, a general construction as described here inevitably depends on multiple derivatives of the solution, and probably does so in a complicated way.

Three mechanisms were used to obtain (14). First, when $\mu > 0$, all modes are damped, and those with large wavenumber $|\mathbf{k}|$ are more strongly so. This gives us the spatial regularity that allows us to control the high-wavenumber modes. Second, since we assumed in (13) that the forcing is slow (or at least the fast part is weak), direct forcing on modes having high frequencies is weak because of frequency mismatch. As $\varepsilon \to 0$, the frequencies of the ageostrophic modes grow larger so the direct forcing on them grows weaker, while the damping remains the same. Ignoring the nonlinear terms, these two mechanisms imply that the ageostrophic energy would decay in time to an $O(\varepsilon)$ quantity at which the forcing is balanced by damping. Third and finally, nonlinear interactions were handled using the well-known fact that there is no fast–fast–slow resonance in the PE, extended to account for the fact that the PDE has triples that are arbitrarily close to resonances (by proving that the strength of near-resonant nonlinear interactions decay sufficiently rapidly in $|\mathbf{k}|$). The rest of the proof is necessary mathematical "details," made possible by the recently obtained regularity results.

[7] While correct for all ε, the bound (14) is only useful when ε is sufficiently small that the rhs is less than the total energy.

We note that the bound (9) on the total energy E_0 was used to obtain the bound (14) on the ageostrophic energy. In other words, we cannot say if arbitrarily large solutions will be nearly quasigeostrophic, but we do know that, for large enough t, our solution will be bounded and close to geostrophic balance. This point will arise again below when we discuss higher-order balance.

Although the end result is qualitatively correct, one may question if our mathematical analysis mirrors (what is believed to be) the physical picture of geostrophic adjustment. Even with the reasonable assumption that μ represents some sort of ''eddy'' rather than ''molecular'' viscosity, one may object that relying completely on viscous damping to obtain geostrophy is, at best, too heavy-handed. We believe that at least two other physical processes, whose rigorous mathematical analyses remain challenging problems, may contribute to the attainment of balance: gravity wave radiation and turbulent transfer. For the first, following the classical picture in Gill (1982), one would like to send fast gravity waves away (to the ''mesosphere'' where they can be damped without directly affecting the nearly geostrophic flow); for recent progress on this front, see Zeitlin (2008) and references therein. As for the second, progress is unlikely to come before the conceptually simpler Navier–Stokes case [see Foias et al. (2001) and Robinson (2007) for a review] is resolved; we believe that much work remains to be done for the latter.

4. Balance and slow manifolds

For higher-order extensions of geostrophic balance, let us start with a dynamical system of the form

$$\begin{aligned} \frac{dp}{dt} + \frac{1}{\varepsilon} Lp &= F(p,q), \\ \frac{dq}{dt} &= G(p,q), \end{aligned} \tag{15}$$

where L is an invertible antisymmetric linear operator (i.e., its eigenvalues are all imaginary).[8] If we drop the nonlinear rhs, q will remain constant while p will undergo fast oscillations. We therefore call p the fast variable and q the slow variable. The system may be infinite-dimensional, in which case F and G will be nonlinear operators, but for now it is simplest to think of (15) as ODEs in $\mathbb{R}^{m+n}$.

As noted in the last section, in the primitive equations (1) the slow variable q is the geostrophic $(\mathbf{v}^0, b^0)$ and the fast variable p is the ageostrophic $(\mathbf{v}^\varepsilon, b^\varepsilon)$; one can rewrite the PEs (1) in fast and slow variables (see, e.g., Temam and Wirosoetisno 2007), but this is not needed for what follows. In terms of (15), geostrophic balance $(\mathbf{v}^\varepsilon, b^\varepsilon) = 0$ corresponds to the leading-order slow manifold $\{p = 0\}$. The result of the last section can then be restated as: for $t \ge T_g$, one has

$$\|p(t)\|^2 \le \varepsilon C, \tag{16}$$

for some suitable norm $\|\cdot\|$ when the system is infinite-dimensional. Of course this will not be true for any F and G, but the generic form (15) makes the following discussion cleaner.

Let us now see if this can be improved. More precisely, we look for a $\Phi(q; \varepsilon)$ such that

$$\|p(t) - \Phi(q(t); \varepsilon)\|^2 \le o(\varepsilon), \tag{17}$$

for some possibly large time $t \ge T_{**}$. One might want to replace $p - \Phi(q)$ by some $\Xi(p, q)$, but since we are looking for higher-order extension of geostrophic balance $p = 0$, the implicit function theorem suggests that the simpler form suffices for suitably bounded solutions. We also require that, at the formal level, our construction is sufficiently general, in that it does not rely on special structures of the nonlinearity (although convergence proofs for PDEs will depend on the nonlinearity). For a survey of various approaches on slow manifolds, see MacKay (2004).

Following Lorenz (1986), we look for a manifold of the form

$$p = \Phi(q; \varepsilon) \tag{18}$$

on which the solution remains if it is initially there. One says that the fast variable p is ''slaved'' to the slow variable q, its time derivative[9] being given by

$$\frac{dp}{dt} = (D\Phi)(q; \varepsilon) \cdot \frac{dq}{dt} = (D\Phi)(q; \varepsilon) \cdot G(\Phi(q; \varepsilon), q), \tag{19}$$

which is completely slow since it does not contain terms of order $1/\varepsilon$ (here $D\Phi$ is the linearization of Φ). The

[8] While a spectral gap (present in the PE) is helpful, it is not absolutely necessary for our construction; that is, the eigenvalues of L may accumulate at 0 (see, e.g., Gallagher and Saint-Raymond 2007).

[9] We note that so far we have loosely identified gravity waves as fast and vortical motion (with nonzero potential vorticity) as slow. There are instances where this identification fails: D. J. Muraki (2003, unpublished manuscript) gave an example of exponentially small stationary gravity waves in the lee of a mountain. Such stationary gravity waves will not be captured by our construction. We thank J. Vanneste for pointing this out.

problem now is to find Φ, which must satisfy the functional equation

$$(D\Phi)\cdot G(\Phi,q)+\frac{1}{\varepsilon}L\Phi-F(\Phi,q)=0. \quad (20)$$

Various methods may be used to construct Φ perturbatively in ε. One can expand Φ as a power series in ε (see, e.g., Warn et al. 1995), but the mathematical analysis is simpler if we use the following iterative scheme. Let $\Phi^0(q;\varepsilon)=0$ and, for $n=1,\ldots,$

$$\frac{1}{\varepsilon}L\Phi^{n+1}(q;\varepsilon)=F(\Phi^n(q;\varepsilon),q) - D\Phi^n(q;\varepsilon)\cdot G(\Phi^n(q;\varepsilon),q). \quad (21)$$

In general, this process is asymptotic rather than convergent, as the repeated differentiations in $D\Phi^n$ eventually cause a loss of analyticity.[10] For ODEs with F and G analytic, Cotter (2004) proved that it is possible to carry out the iteration (21) to $n_*\sim 1/\varepsilon$, resulting in a $\Phi^*:=\Phi^{n*}$ on which the normal velocity is of order $\exp(-c/\varepsilon)$ for q in a compact set M. This shows that as long as q remains in M, the manifold $\Phi^*(q;\varepsilon)$ is invariant up to an exponentially small error. Note that this does not imply that the solution stays near Φ^* for exponentially long times.

It is known that in general this result cannot be improved qualitatively: there are examples (see, e.g., Kruskal and Segur 1991; Vanneste 2008) for which (20) has no solution and (21) eventually diverges.

Formally, the iteration (21) gives us a series of manifolds Φ^n, each slower and more invariant than the previous one, but a more careful inspection tells us that for fixed q and ε the radius of analyticity of Φ^n shrinks to zero as $n\to\infty$. This is the reason why the exponentially slow manifold Φ^* is not defined for all values of q, but only for $q\in M$: once M is fixed, the iteration (21) only produces (strictly speaking, is only guaranteed to produce) a smooth manifold Φ^n for $n\le n_*(\varepsilon; M)$, with larger M generally resulting in smaller n_*. In the infinite-dimensional systems for which this construction is possible, a condition even stronger than compactness is needed.

Once Φ^* is found, one can show that the dynamics on it can be approximated, again up to an exponentially small error, by the "exponential balance model"

$$\frac{dq}{dt}=G(\Phi^*(q;\varepsilon),q), \quad (22)$$

in the sense that the solutions of (15) and of (22), with $p(0)=\Phi^*(q(0);\varepsilon)$ for (15), remain exponentially close to each other over a time scale of order 1. In general, this cannot be improved since even if both solutions stay close to the slow manifold (and q remains in M), which may not be the case, they may still diverge exponentially.

Generalizing the above construction to PDEs is a difficult task and presently can be done only on a case-by-case basis. For the primitive equations (1), this has been done in TW10: Let $q:=\nabla^\perp\cdot\mathbf{v}+\partial_z b$ be the quasigeostrophic potential vorticity and assume that the forcing $\mathbf{S}$ is sufficiently smooth (Gevrey) and, unlike at leading order above, time independent. Then for ε sufficiently small (depending on the forcing $\mathbf{S}$, the viscosity μ, σ below and the domain size):[11]

(i) one can find a manifold $\{(\mathbf{v}^\varepsilon, b^\varepsilon)=\Phi^*(q;\varepsilon)\}$ that satisfies the analog of (20) up to an exponentially small error [i.e., instead of 0, the right-hand side of (20) is of order $\exp(-\sigma/\varepsilon^{1/3})$]; and
(ii) our solution will be exponentially close to this slow manifold after some time; that is, for $t\ge T_*$,

$$\int|[\mathbf{v}^\varepsilon(t),b^\varepsilon(t)]-\Phi^*[q(t);\varepsilon]|^2\,d\mathbf{x}^3 \le C_*(\mathbf{S};\mu,\sigma)\exp(-2\sigma/\varepsilon^{1/3}). \quad (23)$$

Here $\sigma>0$ is arbitrarily fixed, but C_* depends on it.

As in the finite-dimensional case above, Φ^* is not defined for all q, but only for those in the compact set[12] $E_1^\sigma\le K_1^\sigma$. The bound (11) is therefore essential in order to prove (23). Not explicitly indicated above is the dependence of Φ^* on $\mathbf{S}$, μ, and the domain size; in addition, the time T_* also depends on the initial conditions $U(0)$.

One may ask why we used the quasigeostrophic potential vorticity q instead of Ertel's potential vorticity $q_E=\nabla^\perp\cdot\mathbf{v}+\partial_z b+\varepsilon[\partial_z b(\nabla^\perp\cdot\mathbf{v})-(\partial_z\mathbf{v})\cdot\nabla^\perp b]$, which is materially conserved exactly in the inviscid case. Although conceptually appealing, the nonlinearity in q_E would make the analysis much more difficult and the fact that q is linear in $(\mathbf{v},b)$ seems to outweigh the fact that it is not an inviscid invariant.

The reader may have noticed that little has been said as to whether Φ^*, or its neighborhood, is slow. The bounds on time derivatives near Φ^* in the proof of (23) tell us that any fast oscillation must be exponentially weak. Therefore, instead of a slow manifold, we have an exponentially thin set (neighborhood of a manifold)—the fuzzy manifold of Lorenz and Warn—which is both exponentially

[10] As pointed out in Warn (1997), this problem is absent for steady solutions.

[11] In TW10, the theorems were stated and proved for $\varepsilon^{1/4}$, but as noted in remark 7 on p. 446, similar computation can be done for $\varepsilon^{1/3}$.

[12] Note that Φ^* is thus a manifold with boundary; this is to be kept in mind in what follows.

slow and forward invariant (meaning that solutions inside this set will remain within it for all $t \geq 0$, but in general not for $t < 0$). We can in fact do better: the attractor $\mathcal{A}$, which is finite dimensional and therefore has zero "thickness," must lie inside this set; $\mathcal{A}$ is invariant[13] by definition and, arguing as above, is slow. If what is known about the Lorenz attractor and other finite-dimensional examples is any guide, $\mathcal{A}$ is (in general) a complicated set that is not a manifold (from a mathematical point of view, this is yet to be proven for our system). It should be noted that $\mathcal{A}$ may not be completely devoid of gravity waves: since our system is forced–dissipative, a solution on $\mathcal{A}$ could be continuously generating (exponentially weak) gravity waves that are continuously damped away.

At the conceptual or physical level, the main ingredients of the proof are the above asymptotic construction plus the three mechanisms used in the last section. Mathematically, there is the added difficulty of finding function spaces to (i) carry out the iteration (21) and (ii) integrate (22). Inspired by Matthies (2001) and ideas from inertial manifolds discussed below, we used the Gevrey regularity of the solution to split the solution spectrally into high- and low-wavenumber components, with the threshold wavenumber κ scaling as $\varepsilon^{-1/3}$. The high component is exponentially small, and the low component is a system of ODEs whose dimension depends on ε. The power $\varepsilon^{1/3}$ in the exponential arose from the fact that we need to bound the spatial derivatives or, equivalently, the κ-dependent constants in the ODEs; we believe that, barring the discovery of some magical cancellations, one cannot do much better than $\varepsilon^{1/3}$. Brushing aside the different models and setups used, another way to narrow the gap between the upper bound of $\exp(-c'/\varepsilon^{1/3})$ in TW10 and the lower bound of approximately $\exp(-c/\varepsilon)$ of Vanneste and Yavneh (2004) and Ólafsdóttir et al. (2008) is the equally formidable task of finding another exact solution emitting stronger gravity waves.

5. Discussion

In the mathematical literature one has the concepts of inertial manifold and approximate inertial manifold; while these concepts have similarities to the slow manifolds of geophysical fluid dynamics, we feel it important also to spell out their differences here. For (approximate) inertial manifolds, one introduces a spectral truncation N and slaves the high modes $U_>$ to the low modes $U_<$ as $U_> = \Upsilon(U_<)$. Here "high" and "low" correspond to the eigenvalues of the Laplacian Δ, in contrast to Φ in (18), which slaves the fast modes to the slow modes, where "fast" and "slow" correspond to the eigenvalues of the antisymmetric operator L. For some systems such as the Cahn–Hilliard and some reaction–diffusion systems, the slaving relation Υ has been proven to be exactly invariant under the dynamics, but the question is still open for the 2D Navier–Stokes equations and our primitive equations. Notwithstanding the existence of (exact) inertial manifold, a family of approximate inertial manifolds can be used to approximate the global attractor $\mathcal{A}$ (Foias et al. 1988; Titi 1990; Debussche and Temam 1994).

While the proofs in TW10 only apply to the forced–dissipative PE, one might speculate that a similar mechanism may be responsible for the persistence of high-order balance observed in numerical simulations (e.g., Mohebalhojeh and McIntyre 2007). It is conceivable that, at least for sufficiently "nice" initial conditions, a tiny amount of viscosity (e.g., in the numerical method) may be sufficient to keep the solution balanced to a high degree over interesting time scales. This was also suggested by MacKay (2004), although no proof is offered.

As noted above, the method used in TW10 uses global norms and no spatially local flow information is used. By taking into account how gravity waves are emitted and dissipated, stronger results could presumably be obtained. Instead of the thin layer around a slow manifold, one would have to consider such objects as the "quasi-manifold" used by Ford et al. (2000) in the small Froude number limit [see Zeitlin (2008), McIntyre (2009), and references therein for recent progress on this front]. Building the analytical machinery to tackle this problem, however, remains a challenge for mathematicians.

Acknowledgments. This research was partially supported by the National Science Foundation under Grant NSF-DMS-0906440 and by the Research Fund of Indiana University. We thank M. E. McIntyre, J. Vanneste, and two anonymous referees for constructive comments and discussion.

REFERENCES

Babin, A., A. Mahalov, and B. Nicolaenko, 2000: Fast singular oscillating limits and global regularity for the 3D primitive equations of geophysics. *Math. Model. Numer. Anal.,* **34,** 201–222.

Baer, F., and J. J. Tribbia, 1977: On complete filtering of gravity modes through nonlinear initialization. *Mon. Wea. Rev.,* **105,** 1536–1539.

Bartello, P., 1995: Geostrophic adjustment and inverse cascades in rotating stratified turbulence. *J. Atmos. Sci.,* **52,** 4410–4428.

Bourgeois, A. J., and J. T. Beale, 1994: Validity of the quasigeostrophic model for large-scale flow in the atmosphere and ocean. *SIAM J. Math. Anal.,* **25,** 1023–1068.

Cao, C., and E. S. Titi, 2007: Global well-posedness of the three-dimensional viscous primitive equations of large-scale ocean and atmosphere dynamics. *Ann. Math.,* **166,** 245–267.

[13] As pointed out above, $\mathcal{A}$ is invariant both forward and backward in time while the thin set containing it is only forward invariant.

Charney, J. G., 1948: On the scale of atmospheric motions. *Geofys. Publ.,* **17,** 3–17.

Cotter, C. J., 2004: Model reduction for shallow water dynamics: Balance, adiabatic invariance and subgrid modelling. Ph.D. thesis, Imperial College, London, 211 pp.

Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.

Debussche, A., and R. Temam, 1994: Convergent families of approximate inertial manifolds. *J. Math. Pure Appl.,* **73,** 489–522.

Errico, R. M., 1984: The statistical equilibrium solution of a primitive-equation model. *Tellus,* **36A,** 42–51.

Foias, C., and R. Temam, 1989: Gevrey class regularity for the solutions of the Navier–Stokes equations. *J. Funct. Anal.,* **87,** 359–369.

——, O. Manley, and R. Temam, 1988: Modelling of the interaction of small and large eddies in two-dimensional turbulent flows. *Math. Model. Numer. Anal.,* **22,** 93–118.

——, O. P. Manley, R. Rosa, and R. Temam, 2001: *Navier–Stokes Equations and Turbulence*. Cambridge University Press, 347 pp.

Ford, R., M. E. McIntyre, and W. A. Norton, 2000: Balance and the slow quasi-manifold: Some explicit results. *J. Atmos. Sci.,* **57,** 1236–1254.

Gallagher, I., and L. Saint-Raymond, 2007: Mathematical study of the betaplane model: Equatorial waves and convergence results. *Mémoires de la Société Mathématique de France,* Vol. 108, 116 pp.

Gill, A. E., 1982: *Atmosphere–Ocean Dynamics*. Academic Press, 662 pp.

Ju, N., 2007: The global attractor for the solutions to the 3D viscous primitive equations. *Discrete Contin. Dyn. Syst.,* **17A,** 159–179.

Kobelkov, G. M., 2007: Existence of a solution ''in the large'' for ocean dynamics equations. *J. Math. Fluid Mech.,* **9,** 588–610.

Kruskal, M. D., and H. Segur, 1991: Asymptotics beyond all orders in a model of crystal growth. *Stud. Appl. Math.,* **85,** 129–181.

Kukavica, I., and M. Ziane, 2007: On the regularity of the primitive equations of the ocean. *Nonlinearity,* **20,** 2739–2753.

Leith, C. E., 1980: Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.,* **37,** 958–968.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.,* **20,** 130–141.

——, 1980: Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci.,* **37,** 1685–1699.

——, 1986: On the existence of a slow manifold. *J. Atmos. Sci.,* **43,** 1547–1557.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model, with applications to normal mode initialization. *Beitr. Phys. Atmos.,* **50,** 253–271.

MacKay, R. S., 2004: Slow manifolds. *Energy Localisation and Transfer,* T. Dauxois et al., Eds., World Scientific, 149–192.

Matthies, K., 2001: Time-averaging under fast periodic forcing of parabolic partial differential equations: Exponential estimates. *J. Differ. Equations,* **174,** 133–180.

McIntyre, M. E., 2009: Spontaneous imbalance and hybrid vortex–gravity structure. *J. Atmos. Sci.,* **66,** 1315–1326.

Mohebalhojeh, A. R., and M. E. McIntyre, 2007: Local mass conservation and velocity splitting in PV-based balanced models. Part II: Numerical results. *J. Atmos. Sci.,* **64,** 1794–1810.

Ólafsdóttir, E. I., A. B. Olde Daalhuis, and J. Vanneste, 2008: Inertia–gravity-wave radiation by a sheared vortex. *J. Fluid Mech.,* **596,** 169–189.

Petcu, M., and D. Wirosoetisno, 2005: Sobolev and Gevrey regularity results for the primitive equations in three space dimensions. *Appl. Anal.,* **84,** 769–788.

Robinson, J. C., 2007: Parametrization of global attractors, experimental observations, and turbulence. *J. Fluid Mech.,* **578,** 495–507.

Temam, R., 1997: *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*. 2nd ed. Springer-Verlag, 648 pp.

——, and D. Wirosoetisno, 2007: Exponential approximations for the primitive equations of the ocean. *Discrete Contin. Dyn. Syst.,* **7B,** 425–440.

——, and ——, 2010: Stability of the slow manifold in the primitive equations. *SIAM J. Math. Anal.,* **42,** 427–458.

Titi, E. S., 1990: On approximate inertial manifolds to the Navier–Stokes equations. *J. Math. Anal. Appl.,* **149,** 540–557.

Tucker, W., 2002: A rigorous ODE solver and Smale's 14th problem. *Found. Comput. Math.,* **2,** 53–117.

Vanneste, J., 2008: Exponential smallness of inertia–gravity wave generation at small Rossby number. *J. Atmos. Sci.,* **65,** 1622–1637.

——, and I. Yavneh, 2004: Exponentially small inertia–gravity waves and the breakdown of quasigeostrophic balance. *J. Atmos. Sci.,* **61,** 211–223.

Vautard, R., and B. Legras, 1986: Invariant manifolds, quasi-geostrophy and initialization. *J. Atmos. Sci.,* **43,** 565–584.

Warn, T., 1997: Nonlinear balance and quasi-geostrophic sets. *Atmos.–Ocean,* **35,** 135–145.

——, and R. Ménard, 1986: Nonlinear balance and gravity-inertial wave saturation in a simple atmospheric model. *Tellus,* **38A,** 285–294.

——, O. Bokhove, T. G. Shepherd, and G. K. Vallis, 1995: Rossby number expansions, slaving principles, and balance dynamics. *Quart. J. Roy. Meteor. Soc.,* **121,** 723–739.

Zeitlin, V., 2008: Decoupling of balanced and unbalanced motions and inertia–gravity wave emission: Small versus large Rossby numbers. *J. Atmos. Sci.,* **65,** 3528–3542.